Handling losses in translog profit models

Bos, Jaap W.B.; Koetter, Michael

Handling losses in translog profit models

Journal:	*Applied Economics*
Manuscript ID:	APE-07-0019.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	12-Aug-2008
Complete List of Authors:	Bos, Jaap; Utrecht University, Utrecht School of Economics Koetter, Michael; University of Groningen, International Economics and Business
JEL Code:	C24 - Truncated and Censored Models < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods, G21 - Banks|Other Depository Institutions|Mortgages < G2 - Financial Institutions and Services < G - Financial Economics
Keywords:	profit efficiency, stochastic frontier analysis, truncation, censoring

Handling losses in translog profit models

J.W.B. Bos [a,*], M. Koetter [b]

[a] Utrecht School of Economics, Utrecht University, Janskerkhof 12, 3512 BL, Utrecht, the Netherlands
[b] University of Groningen, Faculty of Economics, P.O. Box 800, 9700 AV Groningen, the Netherlands

Abstract

In this paper, we compare standard approaches used to handle losses in logarithmic profit models with a simple novel approach. We estimate translog stochastic profit frontiers, and discuss discriminatory power, rank stability and the precision of profit efficiency scores. Contrary to existing methods, our approach does not result in a loss of observations. Our new method enhances rank stability and discriminatory power, and improves the precision of profit efficiency scores.

Key words: profit efficiency, stochastic frontier analysis, truncation and censoring.
JEL: G21, C24

1. Introduction

Profit models play an important role when we assess the determinants of firm profitability or when we benchmark firms' success at maximizing profits. When estimating profit models, we employ (semi-)flexible functional forms like the translog. This is problematic if firms incur losses in our sample, since the logarithm of non-positive numbers is not defined. Hence, we face an important inconsistency between our theoretical model and our empirical specification. As a result, we may lose information on a significant part of our sample. This part, consisting of loss-incurring firms, is often of particular concern, for example when we benchmark firm performance and try to see if poor performance helps predict a firm's exit from the market (see Bos et al., 2008).

⋆ We thank David Humphrey and Clemens Kool for encouraging us to write this paper, and Allen Berger, Jaap Bikker, William Greene, Barbara Meller and Laura Spierdijk for helpful comments. We gratefully acknowledge support from the Deutsche Bundesbank. Michael Koetter gratefully acknowledges support by the Netherlands Organization for Scientific Research (NWO). The usual disclaimer applies.
* Corresponding author.
Email addresses: j.bos@econ.uu.nl (J.W.B. Bos), m.koetter@rug.nl (M. Koetter).

Preprint submitted to Applied Economics *August 12, 2008*

In this paper, we compare standard methods used to handle theses losses with a novel method.[1] We compare both methods by estimating stochastic profit frontiers, where we can compare profit efficiency (PE). When we study profit efficiency scores, we are concerned both with the ability of specifications to discriminate between profit making and loss incurring firms, as well as their capability of achieving rank stability.

Our paper is structured as follows. First, we explain the two most frequently used specifications to handle negative profits as well as our suggested alternative. Then, we introduce our data. Next, we test whether our specification improves the discriminatory power and rank stability of our model. Finally, we conclude.

2. Methodology

Theoretically, firms maximize profits by choosing in- and output quantities at prevailing prices subject to a production technology constraint. Most banking studies employ a modified model by Humphrey and Pulley (1997), that allows for price setting power in output markets.[2] To implement either of the two models empirically, most studies follow Hasenkamp's (1976) early suggestion and use sufficiently flexible functional forms with regard to curvature. The translog functional form proved particularly suited for bank efficiency studies as it is flexible enough to fit the underlying production process and facilitates convergence when maximizing the likelihood function (Berger and Mester, 1997).

We use a true fixed-effect model, where inefficiency scores are i.i.d. and no particular pattern of evolution of inefficiency is specified (see Greene, 2002):

$$\ln \pi_{kt}(w,y,z) = a_k + \sum_{i=1}^{I} a_i \ln m_{ikt} + \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J} a_{ij} \ln m_{ikt} \ln m_{jkt} + \varepsilon_{kt}. \quad (1)$$

Here m consists of outputs y, input prices w, a control variable z (equity), and a time trend t that captures technological change.

We impose the usual homogeneity and symmetry restrictions as in e.g. Lang and Welzel (1996). Whereas the standard profit (and cost) function is the dual to the output distance function that characterizes production technology (i.e. the transformation function), the alternative profit function is the dual to the output distance function and the pricing opportunity set $g(p,y,w)$ (cf. Kumbhakar and Lovell, 2003). The latter "captures the producer's ability to transform exogenous (y,w) into endogenous product prices p" (p. 213). Kumbhakar and Lovell (2003, p. 213) write: "it is reasonable to assume that [the alternative profit function] is nondecreasing in the elements of y and non-increasing in the elements of w." Summing up, whereas imposing homogeneity of degree one on both outputs and input prices is indeed needlessly restrictive, our restrictions on input prices do not violate the approach suggested by Kumbhakar and Lovell (2003).

We assume that ε_{kt} consists of a noise component ν_{kt}, and an inefficiency component u_{kt}, where $\varepsilon_{kt} = \nu_{kt} - u_{kt}$. Here, ν_{kt} is normally distributed, i.i.d. with $\nu_{kt} \sim N(0,\sigma_\nu^2)$. The inefficiency term u_{kt} is drawn from a non-negative half-normal distribution and

[1] For an overview and a solution to this problem for a non-parametric (DEA) model, see Färe et al. (2004).

[2] The alternative profit model specifies an additional constraint: the pricing opportunity set. Banks choose prices for given output quantities subject to this and the technology constraint.

i.i.d. with $u_{kt} \sim |N(0, \sigma_u^2)|$. Point estimates of PE are obtained using the expected value of u_{kt} given ε_{kt} (Jondrow et. al, 1982).

The problem that we address in this paper arises because we assume in our theoretical model that $\pi \epsilon \mathbb{R}$, whereas in our empirical specification $\ln(\pi)$ is not defined if $\pi \epsilon \mathbb{R}_{-}$, where $\pi = [0, -\infty\rangle$. In the literature, we find two dominant solutions to this problem, listed as [i] and [ii] in Table 1. [3]

First, we can truncate π and estimate our model only for those firms where $\pi \epsilon \mathbb{R}_{+}^{*}$, since π is then $\langle 0, \infty\rangle$. In our view, this approach suffers from two shortcomings: (a) it prohibits us from obtaining efficiency scores for loss incurring firms, and (b) not adjusting for truncation leads to biased results (for ordinary least squares (OLS) estimators, see Greene, 2003, Chapter 20). Second, we can rescale π, to ensure that $\pi \epsilon \mathbb{R}_{+}^{*}$ for all firms, for example by adding the maximum loss observed in the sample plus a small number (usually one) to each π. This appears to be the most popular solution in the literature (cf. Berger and Mester, 1997, Vander Vennet, 2002, Maudos et. al, 2002, Kasman and Yildirim, 2006). However, we cannot control for the effect that this manipulation may have on our error term structure. This is particularly problematic in a stochastic frontier analysis, where we are interested in the composition of total error, rather than coefficient estimates or marginal effects.

Table 1
Specifications

Specification	Left-hand side adjustment (π)		Right-hand side adjustment (NPI)	
	$\pi \epsilon \mathbb{R}_{+}^{*}$	$\pi \epsilon \mathbb{R}_{-}$	$\pi \epsilon \mathbb{R}_{+}^{*}$	$\pi \epsilon \mathbb{R}_{-}$
[i] Truncated	π	exclude	–	–
[ii] Rescaled	$\pi + \lvert\min(\pi^{-})\rvert + 1$	$\pi + \lvert\min(\pi^{-})\rvert + 1$	–	–
[iii] Indicator	π	1	1	$\lvert\pi^{-}\rvert$

Summing up, these approaches either (i) result in a loss of crucial observations, or (ii) they neglect the available information about the truncated part of the distribution of the dependent variable $\ln \pi$. We therefore propose an alternative solution, that is in fact similar to censoring and attempts to make use of the available information on the censored part of π. We also left-censor π, but assign a value of one to those banks with $\pi \epsilon \mathbb{R}_{-}$. We aim to include all information available on the censored part of π and to this end specify an additional independent variable NPI (for Negative Profit Indicator). Consequently, we define NPI to be equal to one for observations where $\pi \epsilon \mathbb{R}_{+}^{*}$ and equal to the absolute value of π for a loss incurring bank. We expect and find a negative coefficient for this variable. Table 1 summarizes the resulting three specifications, including our "Indicator" approach. [4]

[3] Other solutions include of course the use of a so-called distribution free approach, as in e.g. Fernandez de Guevara and Maudos (2002). Also, some earlier studies aggregate firm-level data, before estimating a profit frontier, e.g. Maudos and Pastor (2001). Nonparametric linear programming techniques are used in Färe et al. (2004).

[4] As a caveat we point out that we do not aim to combine ML functions derived for (OLS) limited dependent regressions with ML functions derived for SFA with a composed error term. In our view, this would certainly be the econometrically most correct way to tackle the problem of losses in PE research. To our knowledge no such efforts have been undertaken in the econometric literature and we deem the issue out of the scope of our paper.

3. Data

To estimate our alternative profit frontier, we use balance sheet and profit and loss account data for all German banks that reported to the Deutsche Bundesbank between 1993 and 2004.

Table 2
Descriptive statistics

Variable		$\pi \in \mathbb{R}^*_+$		$\pi \in \mathbb{R}_-$	
π [1])	Profit before tax	10.6	(67.4)	-6.9	(41.9)
y_1 [1])	Interbank loans	377.2	(4,364.7)	648.7	(7,249.6)
y_2 [1])	Customer loans	753.0	(6,724.4)	967.8	(13,683.0)
y_3 [1])	Securities	357.0	(3,635.0)	783.0	(10,789.3)
w_1 [2])	Price of fixed assets	21.8	(454.3)	91.6	(963.9)
w_2 [3])	Price of labor	51.2	(152.7)	64.0	(36.3)
w_3 [2])	Price of borrowed funds	3.9	(25.5)	5.7	(24.1)
z [1])	Equity	57.8	(498.9)	87.3	(809.7)
N	Observations	33,533		658	

Means (standard deviations); [1]) In millions of Euros; [2]) In percentages; [3]) In thousands of Euros; w_1 is depreciation and other expenditures on fixed assets/fixed assets; w_2 is personnel expenses/number of full-time equivalent employees; w_3 is interest expenses/total borrowed funds.

We follow the intermediation approach and report our descriptive statistics for profits, input prices, output quantities, and equity in Table 2. In our sample, around 2% of observations (658) exhibit losses. Although our approach can also be used for firms with zero profits, there are no such firms in our data set.

4. Results

We start by comparing the efficiency distributions from all specifications. Figure 1 shows kernel density plots. We observe that the rescaled specification yields a distribution of PE scores that exhibits the lowest standard deviation and is located the closest to full efficiency. Since the maximum loss in the sample equals 989 million Euros, the impact of rescaling the dependent variable for all banks (averaging 10.2 million Euros) appears to be substantial. However, the high density may largely be due to the inability of unadjusted output quantity and input price variables to explain these profits and, more importantly, discriminate between production plan choices of banks.

However, it is important to note that we have no baseline, 'correct' specification. Put differently, we have to accept the fact that PE scores cannot be validated when comparing our specifications and drawing conclusions. In our comparisons, we test two hypotheses which we reflect what we expect from a 'good' specification:

Hypothesis 1
The efficiency levels of firms with positive profits are on average higher than those of firms with negative profits.

This hypothesis is easily explained, since profits are *maximized* and profit efficiency should therefore - ceteris paribus - be increasing in profit.

Hypothesis 2
The relative efficiency ranking for firms with positive profits is insensitive to the inclusion of firms with negative profits.

This hypothesis merits somewhat more explanation. As discussed above, rescaling profits before taking logs is a so-called non-neutral transformation. The relative distance between profits (*pi*) has been changed. In particular, due to the composed error term in stochastic frontier analyses, it is far from obvious that the efficiency ranking of profit making banks ($\pi \epsilon \mathbb{R}^*_+$) is not affected by this transformation.

In sum, we aim at a specification that can both discriminate between firms making a profit and firms incurring a loss and has stable efficiency ranks. Clearly, specification [i] is of little direct use to us, as it has no information on loss incurring firms.

Figure 1. Kernel density of mean *PE* per specification

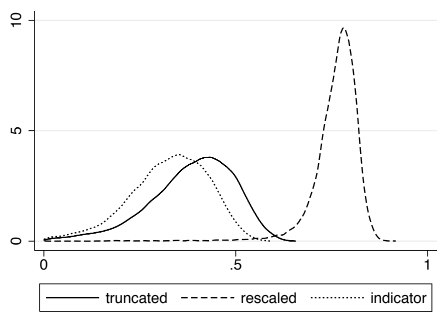


Table 3 therefore lists our comparative statistics for specifications [ii] and [iii]. In both specifications, outliers can influence the overall distribution of efficiency. In fact, we observe that mean efficiency is lower for our indicator approach than for the rescaled specification. To see whether outliers explain this difference, we also present bootstrapped results in Table 4. More importantly, we observe that mean *PE* scores are always higher for profit making firms than for loss incurring firms. These differences are statistically significant, both with and without assuming equal variances. Hence both specifications appear to have sufficient discriminatory power.

Our second hypothesis concerns the ability of specifications to rank profit making firms' efficiency in a stable manner. As several studies have shown, the ability of stochastic frontier models to yield stable ranks is very important (e.g. Bauer et. al, 1998). We calculate ranks for banks with $\pi \epsilon \mathbb{R}^+$ only as our prime interest is the stability of ranks across specifications. Note that the scatterplots are for a comparison vis-à-vis the truncated specification. Also, note that applying truncation to our Indicator approach results in the truncated specification.

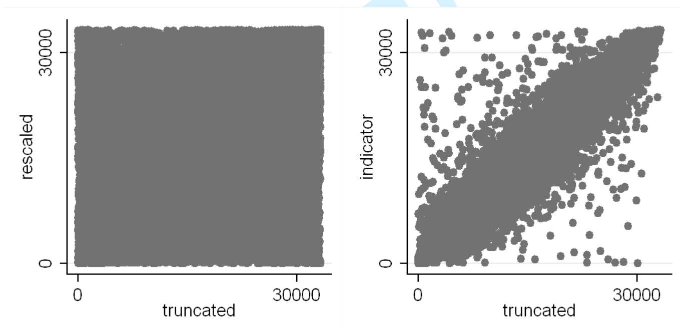
Table 3
Comparative statistics for non-truncated specifications

	[ii] Rescaled	[iii] Indicator
PE for $\pi \in \mathbb{R}^+$	0.755 (0.062)	0.326 (0.102)
PE for $\pi \in \mathbb{R}^-$	0.670 (0.141)	0.175 (0.162)
Independent sample test, equal variances	33.16***	37.19***
Independent sample test, no equal variances	15.35***	23.74***
ρ	0.4698***	0.9717***
KW	375.5***	584.4***

N = 33,533 (π^+), 658 (π^-); IST = Independent samples test, with (1) and without (2) equal variance assumption; ρ = Spearman rank correlations; KW = Kruskal Wallis chi-squared; *** denote significant at the 1% level;. Piecewise correlation between truncated and indicator ranks is 0.971 and significant at the 1% level.

From the plots in Figure 2, we observe that ranks change significantly with the rescaled specification. Many banks are ranked markedly different by the truncated and rescaled specification, respectively. In contrast, our indicator specification ranks much more consistently. The Spearman rank order correlation ρ with the truncated specification is 0.97 and significantly different from zero. Finally, we also perform Kruskal Wallis rank tests, which confirm that profit-making and loss-incurring banks are ranked significantly different from each other. Note, that our indicator specification exhibits substantially higher chi-squared values compared to the rescaled specification.

Figure 2. Correlation



(a) Rescaled (b) Indicator

In addition to testing whether estimated efficiency distributions are identical, we also conduct a bootstrap analysis along the lines of Atkinson and Wilson (1995) to obtain standard errors of mean efficiency estimates (cf. Koetter, 2006). Thereby we can test the precision of PE estimates obtained with the three alternative approaches to handle negative profits, respectively. We follow their suggestion and draw $j = 1, .., J$ bootstrap samples with replacement of the original size N, i.e. 34,191 observations, where $J \simeq 1000$. For each draw j, we estimate mean PE_j^* for the three approaches, respectively. For specification [i], the truncation approach, we obviously only sample 33,533 observations. We denote the mean statistic obtained with the original sample as PE_{obs} and calculate standard errors.[5]

Table 4
Bootstrapped standard errors of mean PE

Specification	PE_{obs} 1)	$\overline{PE}^{*}$ 2)	SE 3)	LB 4)	UB 5)
Truncated	39.3	38.1	0.24	38.9	39.8
Rescaled	75.3	81.1	4.65	66.2	84.4
Indicator	32.3	32.3	0.27	31.8	32.8

Notes: Bootstrapping results for 1,000 repetitions of full resampling with replacement; 1) Mean PE with original full sample; 2) Average mean efficiency after bootstrapping; 3) Standard errors; 4) Lower bound; 5) Upper bound.

In Table 4 we report bootstrapped standard errors and according confidence intervals at the five percent level for all profit models, respectively.[6] Bootstrapped standard errors are largest for the rescaled model. In contrast, the precision of efficiency estimates obtained from our indicator approach closely resembles that obtained for the case when loss-making banks are excluded from the sample. Hence, with our approach we gain discriminatory power without a loss of precision and possible outliers do not bias (mean) efficiency.

5. Conclusion

In this paper, we compare a novel approach to handling losses in translog profit models with specifications that rely on truncation [i] or rescaling [ii] of the dependent variable. We study the effect on stochastic frontier profit efficiency scores. The latter specifications either do not yield any efficiency scores for loss incurring firms [i], or yield efficiency scores which are discriminatory but not stable [ii]. Censoring is shown to greatly improve the rank stability of efficiency scores. In addition, our indicator specification improves the discriminatory power of our translog profit model. Finally, bootstrapped standard errors show that the precision of the indicator approach is substantially higher than for the certainty of efficiency estimates obtained after scaling up all data.

References

Atkinson, S. E. and P. W. Wilson (1995), "Comparing Mean Efficiency and Productivity Scores from Small Samples: A Bootstrap Methodology," *Journal of Productivity Analysis*, 7, 257–282.

Bauer, P. W., A. N. Berger, G. D. Ferrier, and D. B. Humphrey (1998), "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods," *Journal of Economics and Business*, 50, 85—114.

Berger, A. N. and L. J. Mester (1997), "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions," *Journal of Banking & Finance*, 21, 895—947.

[5] $\widehat{SE} = \{\frac{1}{J-1}\sum_{j=1}^{J}(PE_j^* - \overline{PE}^*)^2\}^{1/2}$, where $\overline{PE}^* = \frac{1}{J}\sum_{j=1}^{J} CE_j^*$.

[6] Confidence intervals are $[PE_{obs} - t_{1-\alpha/2,k-1}\widehat{SE}, PE_{obs} + t_{1-\alpha/2,k-1}\widehat{SE}]$, where $t_{1-\alpha/2,k-1}$ is the $(1-\alpha/2)^{th}$ quantile of the t-distribution with $k-1$ degrees of freedom.

Bos, J.W.B., M. Koetter, J.W. Kolari and C.J.M. Kool (2008, forthcoming), "Effects of heterogeneity on bank efficiency scores, " *European Journal of Operational Research*, available at doi:10.1016/j.ejor.2008.01.019.

Färe, R., S. Grosskopf, and W.L. Weber (2004), "The Effect of Risk-Based Capital Requirements on Profit Efficiency in Banking Fare, " *Applied Economics*, 36 (15), 1731-1743.

Fernandez de Guevara, J. and J. Maudos (2002), "Inequalities in the Efficiency of the Banking Sectors of the European Union, " *Applied Economics Letters*, 9 (8), 541-544.

Greene, W. H. (2002), *Limdep 8.0 Econometric Modelling Guide*, Volume 2, Econometric Software, Castle Hill, Australia.

Greene, W. H. (2003), *Econometric Analysis*, 5th Edition, Prentice Hall, New Jersey, U.S..

Hasenkamp, G. (1976), *Specification and Estimation of Multiple-Output Production Functions*, Springer, New York, U.S..

Humphrey, D. H. and B. Pulley, Lawrence (1997), "Bank's Response to Deregulation: Profits, Technology and Efficiency," *Journal of Money, Credit, and Banking* 29, 73—93.

Jondrow, J., C. A. K. Lovell, S. Van Materov, and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 19, 233—238.

Kasman, A. and C. Yildirim (2006), "Cost and Profit Efficiencies in Transition Banking: The Case of New EU Members, " *Applied Economics*, 38 (9), 1079-1090.

Koetter, M. (2006), "Measurement Matters - Alternative Input Price Proxies for Bank Efficiency Analyses," *Journal of Financial Services Research*, forthcoming.

Kumbhakar, S.C. and C.A. Knox Lovell (2003), *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge.

Lang, G. and P. Wetzel (1996), "Efficiency and Technical Progress in Banking: Empirical Results for a Panel of German Cooperative Banks, " *Journal of Banking and Finance*, 21, 1003-1023.

Maudos, J. and J.M. Pastor (2001), "Cost and Profit Efficiency in Banking: An International Comparison of Europe, Japan and the USA, " *Applied Economics Letters*, 8 (6), 383-387.

Maudos, J., J.M. Pastor, F. Pérez, and J. Quesada (2002), "Cost and Profit Efficiency in European Banks," *Journal of International Financial Markets, Institutions and Money*, 12, 33-58.

Vander Vennet, R. (2002), "Cost and Profit Efficiency of Financial Conglomerates and Universal Banks in Europe," *Journal of Money, Credit, and Banking*, 34, 254-282.